



US009407540B2

(12) **United States Patent**
Kumar et al.

(10) **Patent No.:** **US 9,407,540 B2**
(45) **Date of Patent:** **Aug. 2, 2016**

(54) **DISTRIBUTED SERVICE CHAINING IN A NETWORK ENVIRONMENT**

(71) Applicant: **CISCO TECHNOLOGY, INC.**, San Jose, CA (US)

(72) Inventors: **Surendra M. Kumar**, San Ramon, CA (US); **Dileep K. Devireddy**, San Jose, CA (US); **Nagaraj A. Bagepalli**, Fremont, CA (US); **Abhijit Patra**, Saratoga, CA (US)

(73) Assignee: **CISCO TECHNOLOGY, INC.**, San Jose, CA (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 312 days.

(21) Appl. No.: **14/020,649**

(22) Filed: **Sep. 6, 2013**

(65) **Prior Publication Data**

US 2015/0071285 A1 Mar. 12, 2015

(51) **Int. Cl.**
G06F 15/173 (2006.01)
H04L 12/725 (2013.01)
H04L 12/741 (2013.01)

(52) **U.S. Cl.**
CPC **H04L 45/306** (2013.01); **H04L 45/54** (2013.01)

(58) **Field of Classification Search**
CPC **H04L 45/306**
USPC **709/238**
See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

4,769,811 A 9/1988 Eckberg et al.

7,644,182 B2 1/2010 Banerjee
7,738,469 B1 6/2010 Shekolkar et al.
2004/0109412 A1 6/2004 Hansson
2005/0204042 A1 9/2005 Banerjee
2005/0289244 A1 12/2005 Sahu et al.
2008/0177896 A1 7/2008 Quinn et al.
2008/0181118 A1 7/2008 Sharma
2008/0209039 A1 8/2008 Tracey et al.
2009/0003364 A1 1/2009 Fendick et al.

(Continued)

FOREIGN PATENT DOCUMENTS

WO WO2015/073190 5/2015

OTHER PUBLICATIONS

USPTO Jun. 18, 2015 Non-Final Office Action from U.S. Appl. No. 14/037,210.

(Continued)

Primary Examiner — Ario Etienne

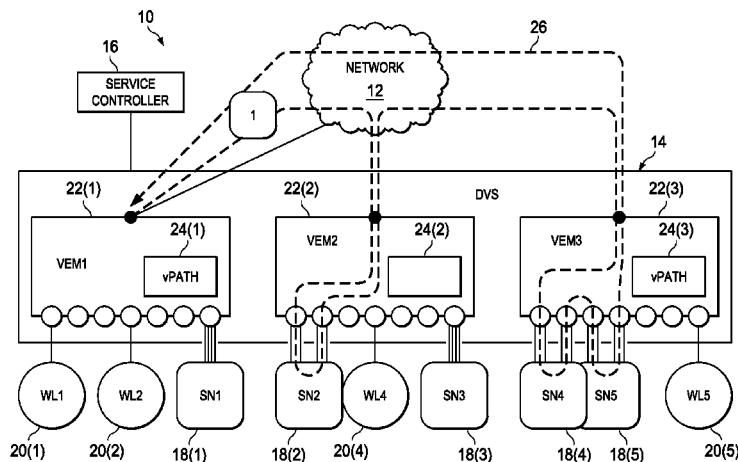
Assistant Examiner — Sahera Halim

(74) *Attorney, Agent, or Firm* — Patent Capital Group

(57) **ABSTRACT**

An example method for distributed service chaining in a network environment is provided and includes receiving a packet belonging to a service chain in a distributed virtual switch (DVS) network environment, wherein the packet includes a network service header (NSH) indicating a service path identifier identifying the service chain and a location of the packet on the service chain, evaluating a service forwarding table to determine a next service node based on the service path identifier and the location, with a plurality of different forwarding tables distributed across the DVS at a corresponding plurality of virtual Ethernet Modules (VEMs) associated with respective service nodes in the service chain, and forwarding the packet to the next service node, with substantially all services in the service chain provided sequentially to the packet in a single service loop on a service overlay.

20 Claims, 7 Drawing Sheets



(56)

References Cited**U.S. PATENT DOCUMENTS**

2011/0023090	A1	1/2011	Asati et al.
2012/0131662	A1	5/2012	Kuik
2013/0003735	A1	1/2013	Chao et al.
2013/0124708	A1	5/2013	Lee
2013/0163594	A1	6/2013	Sharma
2013/0163606	A1	6/2013	Bagepalli et al.
2013/0272305	A1	10/2013	Lefebvre et al.
2013/0311675	A1	11/2013	Kancherla
2014/0036730	A1	2/2014	Nellikar
2014/0115062	A1	4/2014	Liu
2014/0254603	A1	9/2014	Banavalikar
2014/0279863	A1	9/2014	Krishnamurthy
2014/0280836	A1	9/2014	Kumar et al.
2014/0321459	A1	10/2014	Kumar et al.
2015/0052516	A1	2/2015	French
2015/0074276	A1	3/2015	Decusatis
2015/0222516	A1	8/2015	Deval

OTHER PUBLICATIONS

Boucadair et al., "Differentiated Service Function Chaining Framework," Network Working Group Internet Draft draft-boucadair-network-function-chaining-03, Aug. 21, 2013.

Quinn, et al., "Network Service Header," Network Working Group Internet Draft draft-quinn-nsh-01, Jul. 12, 2013.

"Digital Program Insertion," from Wikipedia, the free encyclopedia, Jan. 2, 2012; 1 page http://en.wikipedia.org/w/index.php?title=Digital_Program_Insertion&oldid=469076482.

"Dynamic Adaptive Streaming over HTTP," from Wikipedia, the free encyclopedia, Oct. 25, 2012; 3 pages; http://en.wikipedia.org/w/index.php?title=Dynamic_Adaptive_Streaming_over_HTTP&oldid=519749189.

"G-Streamer and in-band Metadata," from RidgeRun Developer Connection, Jun. 19, 2012; 5 pages; https://developer.ridgerun.com/wiki/index.php/GStreamer_and_in-band_metadata.

"ISO/IEC JTC 1/SC 29, Information Technology—Dynamic Adaptive Streaming over HTTP (DASH)—Part 1: Media Presentation Description and Segment Formats," International Standard © ISO/IEC 2012—All Rights Reserved; Jan. 5, 2012; 131 pages.

"M-PEG 2 Transmission," © Dr. Gorry Fairhurst, 9 pages [Published on or about Jan. 12, 2012] <http://www.erg.abdn.ac.uk/future-net/digital-video/mpeg2-trans.html>.

"MPEG Transport Stream," from Wikipedia, the free encyclopedia, Nov. 11, 2012; 7 pages http://en.wikipedia.org/w/index.php?title=MPEG_transport_stream&oldid=522468296.

Wong, F., et al., "SMPTE-TT Embedded in ID3 for HTTP Live Streaming, draft-smppte-id3-http-live-streaming-00," Informational Internet Draft, Jun. 2012; 7 pages <http://tools.ietf.org/html/draft-smppte-id3-http-live-streaming-00>.

PCT Jan. 19, 2015 International Search Report and Written Opinion from International Application Serial No. PCT/US2014/062244.

Kumar, S., et al., "Service Function Path Optimization: draft-kumar-sfc-sfp-optimization-00.txt," Internet Engineering Task Force, IETF; Standard Working Draft, May 10, 2014.

U.S. Appl. No. 14/037,143, filed Sep. 25, 2013, entitled "Path Optimization in Distributed Service Chains in a Network Environment," Inventor(s): Sourabh Suresh Patwardhan, et al.

U.S. Appl. No. 14/037,210, filed Sep. 25, 2013, entitled "Co-Operative Load Sharing and Redundancy in Distributed Service Chains in a Network Environment," Inventor(s): Maithili Narasimha, et al.

U.S. Appl. No. 14/081,772, filed Nov. 15, 2013, entitled "Shortening of Service Paths in Service Chains in a Communications Network," Inventor(s): Surendra M. Kumar, et al.

"ANSI/SCTE 35 2007 Digital Program Insertion Cueing Message for Cable," Engineering Committee, Digital Video Subcommittee, American National Standard, Society of Cable Telecommunications Engineers, © Society of Cable Telecommunications Engineers, Inc. 2007 All Rights Reserved, 140 Philips Road, Exton, PA 19341; 42 pages.

"CEA-708," from Wikipedia, the free encyclopedia, Nov. 15, 2012; 16 pages <http://en.wikipedia.org/w/index.php?title=CEA-708&oldid=523143431>.

U.S. Appl. No. 15/055,691, filed Feb. 29, 2016, entitled "Shortening of Service Paths in Service Chains in a Communications Network," Inventor(s): Surendra M. Kumar, et al.

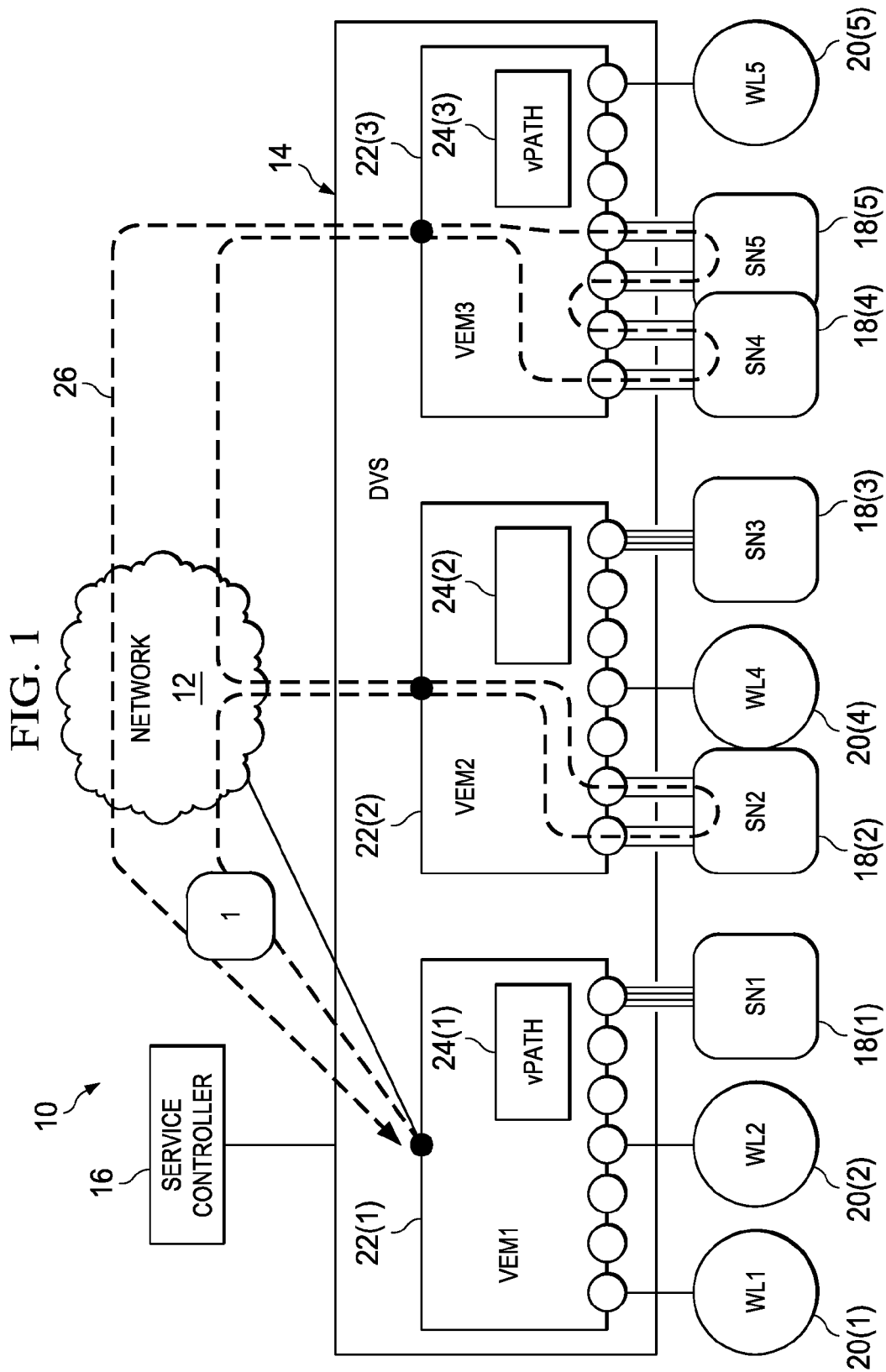
USPTO Sep. 15, 2015 Non-Final Office Action from U.S. Appl. No. 14/037,143.

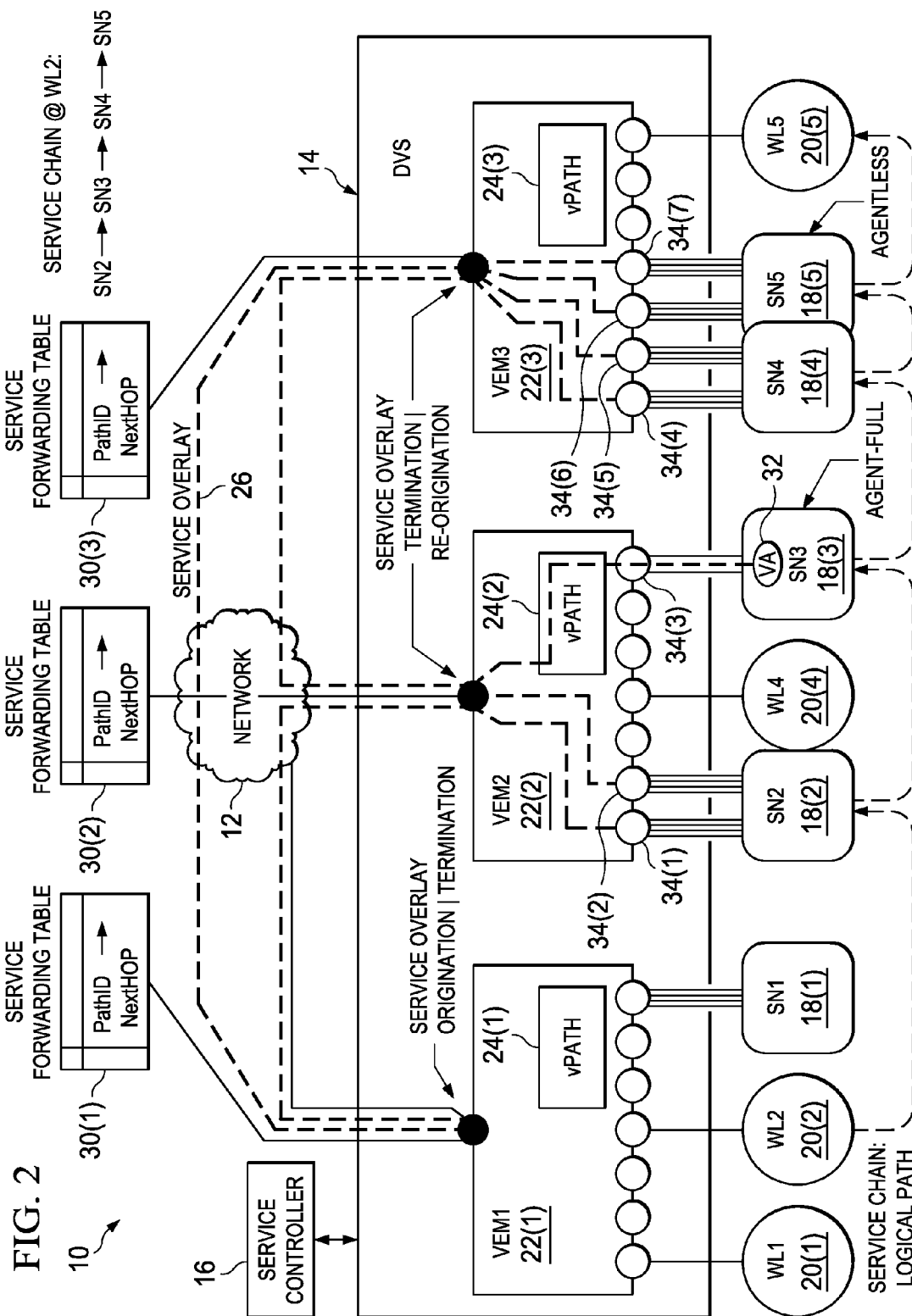
USPTO Nov. 23, 2015 Final Office Action from U.S. Appl. No. 14/037,210.

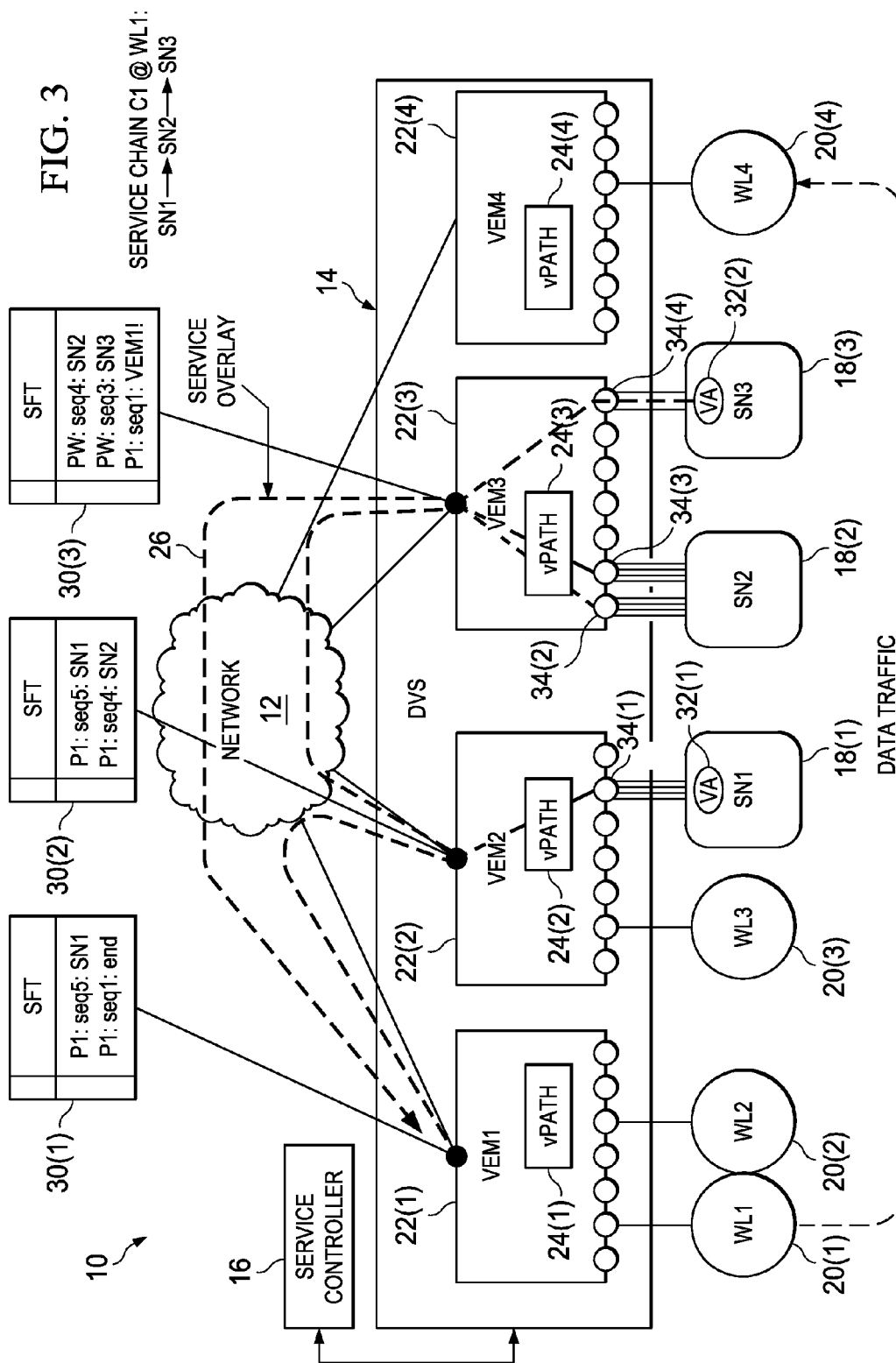
USPTO Aug. 12, 2015 Non-Final Office Action from U.S. Appl. No. 14/081,772.

USPTO Nov. 24, 2015 Notice of Allowance from U.S. Appl. No. 14/081,772.

U.S. Appl. No. 14/997,212, filed Jan. 15, 2016, entitled "Leaking Routes in a Service Chain," Inventor(s).







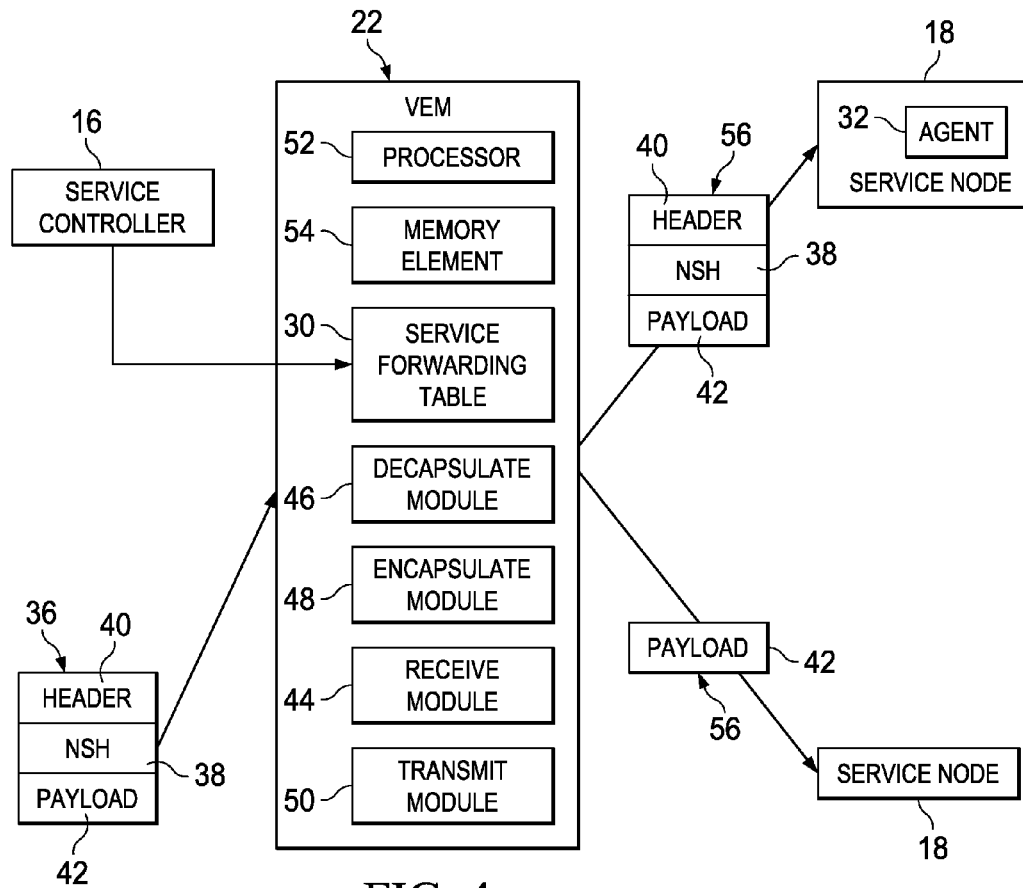


FIG. 4

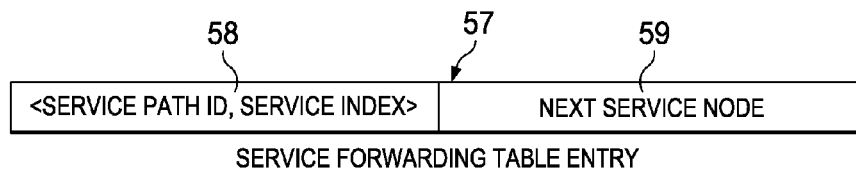
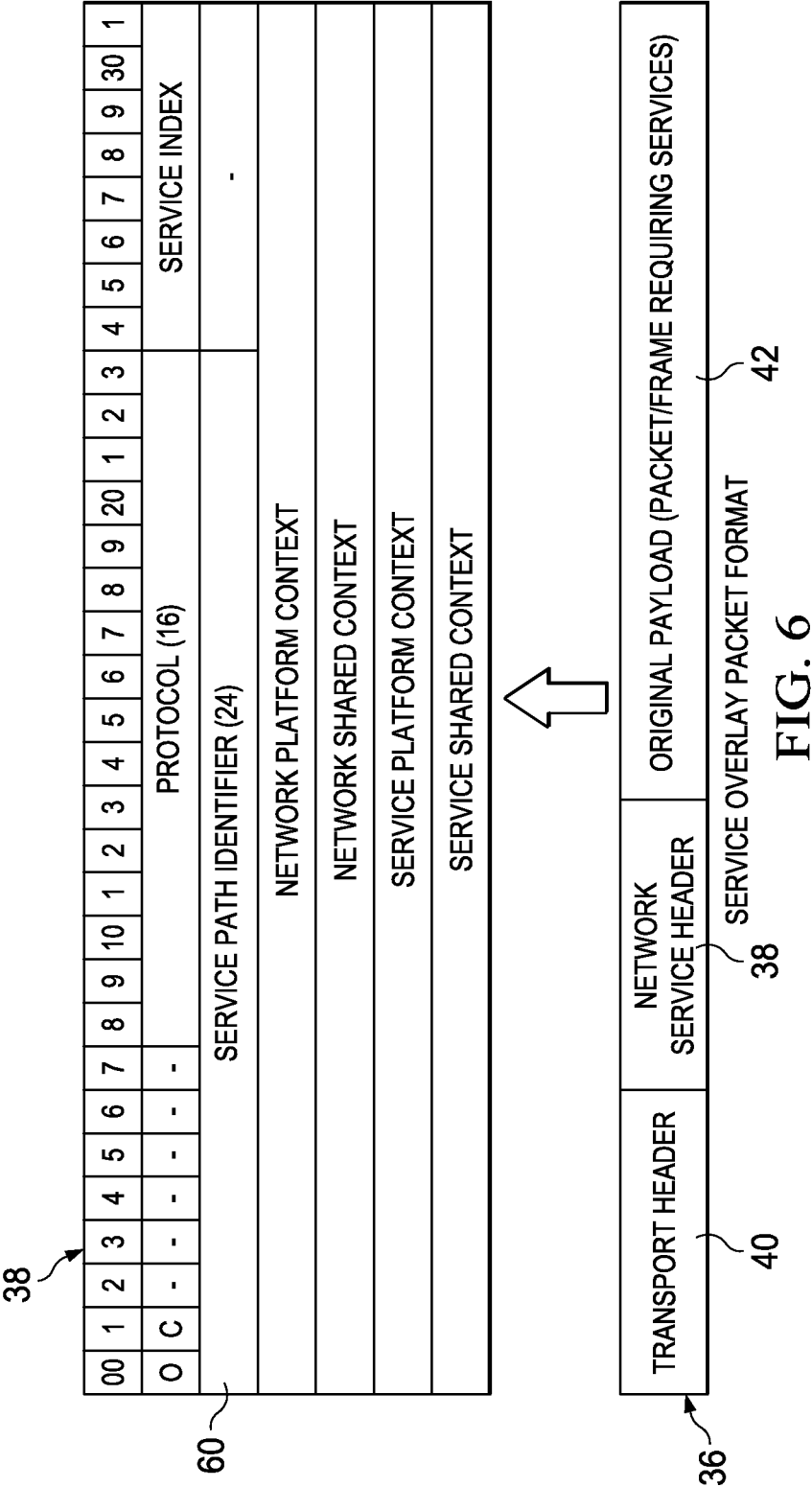


FIG. 5



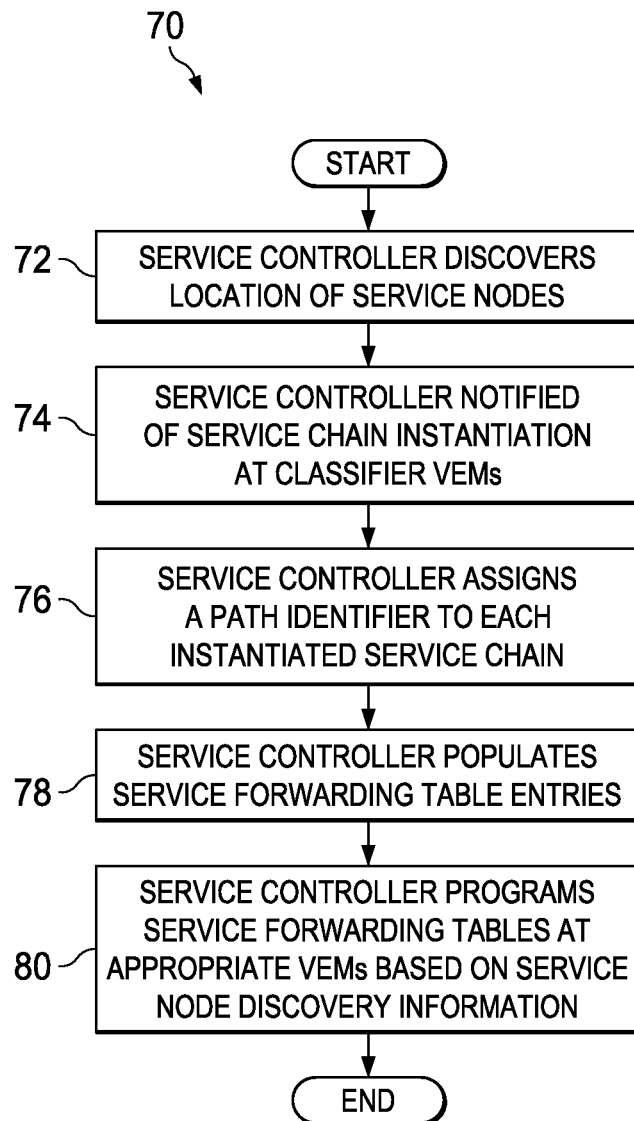
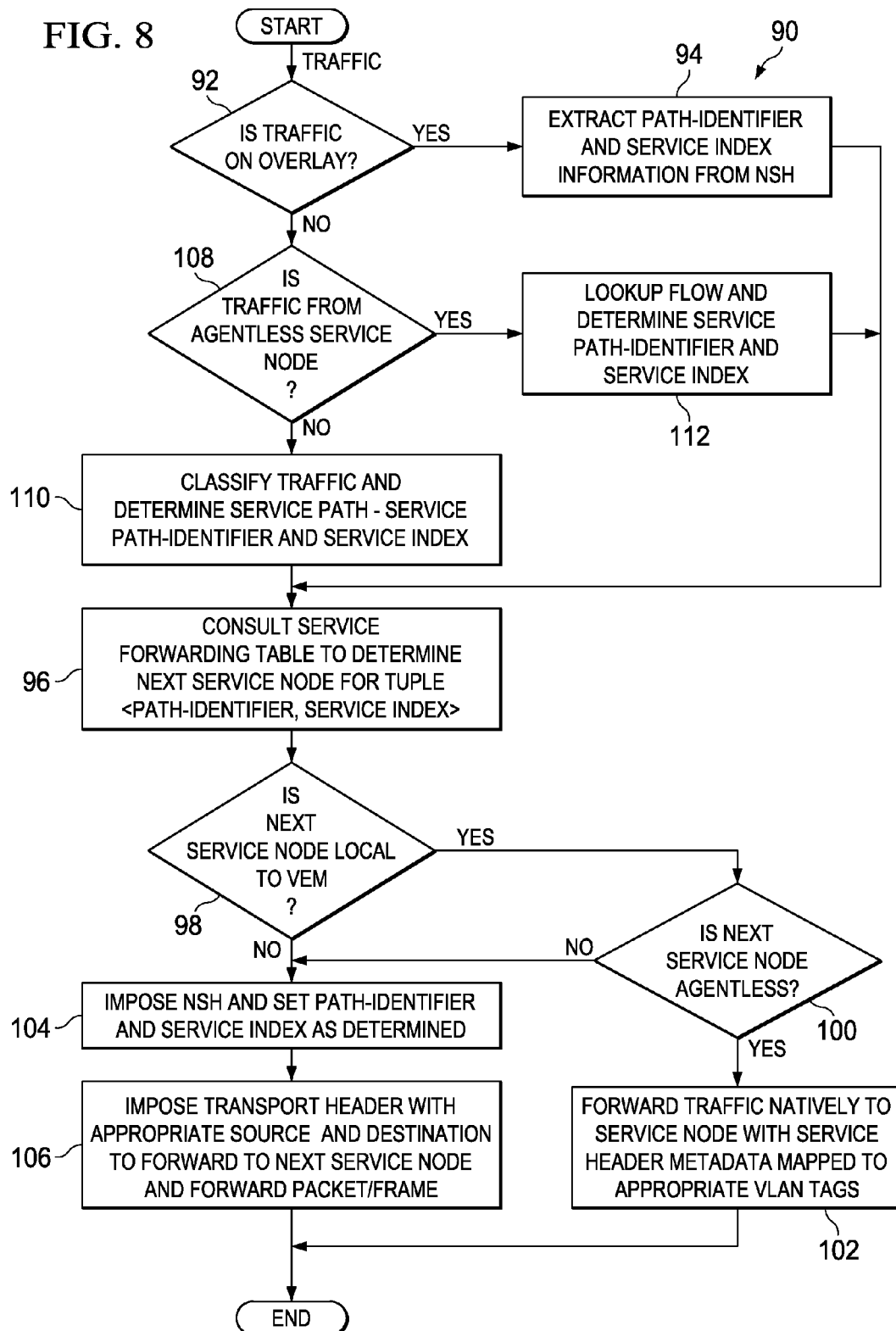


FIG. 7



1

DISTRIBUTED SERVICE CHAINING IN A NETWORK ENVIRONMENT

TECHNICAL FIELD

This disclosure relates in general to the field of communications and, more particularly, to distributed service chaining in a network environment.

BACKGROUND

Data centers are increasingly used by enterprises for effective collaboration and interaction and to store data and resources. A typical data center network contains myriad network elements, including hosts, load balancers, routers, switches, etc. The network connecting the network elements provides secure user access to data center services and an infrastructure for deployment, interconnection, and aggregation of shared resource as required, including applications, hosts, appliances, and storage. Improving operational efficiency and optimizing utilization of resources in data centers are some of the challenges facing data center managers. Data center managers want a resilient infrastructure that consistently supports diverse applications and services and protects the applications and services against disruptions. A properly planned and operating data center network provides application and data integrity and optimizes application availability and performance.

BRIEF DESCRIPTION OF THE DRAWINGS

To provide a more complete understanding of the present disclosure and features and advantages thereof, reference is made to the following description, taken in conjunction with the accompanying figures, wherein like reference numerals represent like parts, in which:

FIG. 1 is a simplified block diagram illustrating a communication system for distributed service chaining in a network environment;

FIG. 2 is a simplified block diagram illustrating example details of an embodiment of the communication system;

FIG. 3 is a simplified block diagram illustrating other example details of an embodiment of the communication system;

FIG. 4 is a simplified block diagram illustrating yet other example details of an embodiment of the communication system;

FIG. 5 is a simplified block diagram illustrating yet other example details of an embodiment of the communication system;

FIG. 6 is a simplified block diagram illustrating yet other example details of an embodiment of the communication system;

FIG. 7 is a simplified flow diagram illustrating example operations that may be associated with an embodiment of the communication system; and

FIG. 8 is a simplified flow diagram illustrating other example operations that may be associated with an embodiment of the communication system.

DETAILED DESCRIPTION OF EXAMPLE EMBODIMENTS

Overview

An example method for distributed service chaining in a network environment is provided and includes receiving a packet belonging to a service chain in a distributed virtual

2

switch (DVS) network environment, where the packet includes a network service header (NSH) indicating a service path identifier identifying the service chain and a location of the packet on the service chain, evaluating a service forwarding table to determine a next service node based on the service path identifier and the location, with a plurality of different forwarding tables distributed across the DVS at a corresponding plurality of virtual Ethernet Modules (VEMs) associated with respective service nodes in the service chain, and forwarding the packet to the next service node, with substantially all services in the service chain provided sequentially to the packet in a single service loop on a service overlay.

Example Embodiments

Turning to FIG. 1, FIG. 1 is a simplified block diagram illustrating a communication system 10 for distributed service chaining in a network environment in accordance with one example embodiment. FIG. 1 illustrates a network 12 (generally indicated by an arrow) comprising a distributed virtual switch (DVS) 14. DVS 14 can include a service controller 16. A plurality of service nodes (SN) 18 (e.g., SNs 18(1)-18(5)) may provide various network services to packets entering or leaving network 12. A plurality of virtual machines (VMs) may provide respective workloads (WLs) 20 (e.g., WL 20(1)-20(5)) on DVS 14, for example, by generating or receiving packets through DVS 14. One or more virtual Ethernet modules (VEMs) 22 (e.g., VEMs 22(1)-22(3)) may facilitate packet forwarding by DVS 14. In various embodiments, DVS 14 may execute in one or more hypervisors in one or more servers (or other computing and networking devices) in network 12. Each hypervisor may be embedded with one of VEMs 22 that can perform various data plane functions such as advanced networking and security, switching between directly attached virtual machines, and uplinking to the rest of the network. Each VEM 22(1)-22(3) may include respective vPaths 24(1)-24(3) that can redirect traffic to SNs 18 before DVS 14 sends the packets into WLs 20.

Note that although only a limited number of SNs 18, WLs 20, VEMs 22, and vPaths 24 are provided in the FIGURE for ease of illustration, any number of service nodes, workloads, VEMs and vPaths may be included in communication system 10 within the broad scope of the embodiments. Moreover, the service nodes and workloads may be distributed within network 12 in any suitable configuration, with various VEMs and vPaths to appropriately steer traffic through DVS 14.

Embodiments of communication system 10 can facilitate distributed service chaining in network 12. As used herein, the term “service chain” includes an ordered sequence of a plurality of services provided by one or more SNs (e.g., applications, virtual machines, network appliances, and other network elements that are configured to provide one or more network services) in the network. A “service” may include a feature that performs packet manipulations over and beyond conventional packet forwarding. Examples of services include encryption, decryption, intrusion management, firewall, load balancing, wide area network (WAN) bandwidth optimization, application acceleration, network based application recognition (NBAR), cloud services routing (CSR), virtual interfaces (VIPs), security gateway (SG), network analysis, etc. The service may be considered an optional function performed in a network that provides connectivity to a network user. The same service may be provided by one or more SNs within the network.

According to some embodiments, a user (e.g., system administrator) can configure the service chain and provision it directly at an applicable workload 20 (e.g., WL 20(1)).

Service controller **16** may segment the user configured service chain in DVS **14**. According to various embodiments, VEMs **22(1)-22(3)** may generate headers for forwarding packets according to the configured service chain such that substantially all services in the service chain may be provided in a single service loop irrespective of the number of services, with respective VEMs **22(1)-22(3)** making independent decisions (e.g., without referring to other VEMs or other network elements) about the next hop decisions in the service chain packet forwarding. As used herein, the term “service loop” refers to a path of the packet from a starting point (e.g., WL **20(1)**) through various service nodes (e.g., SN **18(2)**, SN **18(4)**, SN **18(5)**) of the service chain until termination at the starting point (e.g., WL **20(1)**). The service chain traffic may be steered over network **12** in a service overlay **26**.

As used herein, the term “service controller” includes a process (e.g., instance of a computer program that is executing) that can provision services at one or more service nodes according to preconfigured settings. The preconfigured settings may be provided at the service controller by a user through an appropriate command line interface, graphical user interface, script, or other suitable means. The term “VEM” includes one or more network interfaces, at least some portions of switching hardware and associated firmware and software, and one or more processes managing the one or more network interfaces to facilitate packet switching in a switch, including a distributed virtual switch (e.g., DVS **14**). VEMs may be named as service VEMs when they provide connectivity to service nodes; and as classifier VEMs when they provide connectivity to the workloads that function as the initial node in a service chain.

Service overlay **26** encompasses a level of indirection, or virtualization, allowing a packet (e.g., unit of data communicated in the network) destined to a specific workload to be diverted transparently (e.g., without intervention or knowledge of the workloads) to other service nodes as appropriate. Service overlay **26** includes a logical network built on top of existing network **12** (the underlay). Packets are encapsulated or tunneled to create the overlay network topology. For example, service overlay **26** can include a suitable header (called a network service header (NSH)), with corresponding source and destination addresses relevant to the service nodes in the service chain.

For purposes of illustrating the techniques of communication system **10**, it is important to understand the communications that may be traversing the system shown in FIG. **1**. The following foundational information may be viewed as a basis from which the present disclosure may be properly explained. Such information is offered earnestly for purposes of explanation only and, accordingly, should not be construed in any way to limit the broad scope of the present disclosure and its potential applications.

Service chaining involves steering traffic through multiple services in a specific order. The traffic may be steered through an overlay network, including an encapsulation of the packet to forward it to appropriate service nodes. The services in the chain are typically of two types: agent-full services and agent-less services. Agent-full services host an embedded agent owned by a network infrastructure provider to abstract the underlying network details in inserting services. The embedded agent exposes application programming interfaces (APIs) to the services to enable interaction with the underlying (network) infrastructure for service insertion including communication of metadata to the services and utilizing advanced infrastructure capabilities (e.g., offloads). Agent-less services integrate with the underlying infrastructure in their native forms through respective network interfaces.

Although agent-less services lose the ability to utilize advanced infrastructure capabilities, it eases integration of services. Given the benefits of either type of service, any service chaining architecture should desirably support both types of service in a service chain. The traffic steering mechanism executes either completely in the network infrastructure, or both in the network infrastructure and the agents (that execute in the respective service nodes).

Existing service insertion architectures (such as vPath) support both agent-full and agent-less services. However, typical service chaining includes either agent-full services, or alternatively, agent-less services; the service chains do not include both agent full and agent-less services in the same service chain. In addition, the service chains are orchestrated in a centralized fashion in the network infrastructure. The centralized model of service chaining is termed hub-n-spoke or in-and-out: a network node (switch/router) intercepting and classifying the traffic (requiring services) acts as the hub, while the spokes extend from the central node to the services (e.g., via additional switches and routers on an overlay). Service chaining architecture in such schemes are geared primarily for agent-full services. Further, some service chaining architectures require the services to participate in service forwarding through the embedded agent. A major drawback of the hub-n-spoke service chaining scheme is performance degradation due to the centralized nature of the architecture. Moreover, there does not exist a scheme that chains agent-less and agent-full services while maintaining the service forwarding solely in the network infrastructure.

Communication system **10** is configured to address these issues (and others) in offering a system and method for distributed service chaining in a network environment. Embodiments of communication system **10** may facilitate a distributed method of service chaining that chains both agent-full and agent-less services in the same service chain, without participation of services in service forwarding. Each VEM **22(1)-22(3)** may serve as an originator of respective network service headers (NSHs) for service overlay **26**. As used herein, the term “network service header” includes a data plane header (e.g., metadata) added to frames/packets. The NSH contains information required for service chaining, and metadata added and consumed by SNs **18** and WLs **20**. (Examples of metadata include classification information used for policy enforcement and network context for forwarding post service delivery). According to embodiments of communication system **10**, each NSH may include a service path identifier identifying the service chain to which a packet belongs, and a location of the packet on the service chain, which can indicate the service hop (NSH aware node to forward the packet) on service overlay **26**. The service path identifier and the location of the packet can comprise any suitable text, number or combination thereof. In an example embodiment, the service path identifier is a 24 bit number, and the location may be specified by an 8 bit number.

According to various embodiments, a user may configure (e.g., provision, arrange, organize, construct, etc.) the service chains at service controller **16**. Service controller **16** may discover location of service nodes **18(1)-18(5)**. In some embodiments, the service chain may be provisioned by service controller **16** in a port profile at respective vPaths **24(1)-24(3)** associated with specific workloads **20** that instantiate the service chains, thereby binding the service policy including the service chain with the network policy included in the port profile. In other embodiments, when service chains are instantiated at classifier VEM **22(1)**, associated with the initiating workload **20(2)**, service controller **16** may be notified of the service chain instantiation. Service controller **16** may

assign a path identifier to each instantiated service chain. Service controller 16 may populate service forwarding table entries indicating the next service hop for respective service chains identified by corresponding path identifiers. Service controller 16 may program service forwarding tables at appropriate VEMs 22 based on service node discovery information.

Merely for illustrative purposes, and not as a limitation, assume a service chain 1 provisioned at WL 20(2) as follows: WL 2→SN2→SN4→SN5. In other words, a packet originating at WL 20(2) may be steered to SN 18(2), serviced accordingly, then to SN 18(4), then to SN 18(5), and finally returned to WL 20(2). VEM 22(1) may generate an NSH including the Internet Protocol (IP) or Media Access Control (MAC) address of VEM 22(1) at which WL 20(2) is located as a source address, and an IP/MAC address of SN 18(2) as the next service hop. Destination VEM 22(2), at which SN 18(2) is located may inspect the NSH and take suitable actions depending on whether SN 18(2) includes an agent (e.g., agent-full) or does not include an agent (e.g., agent-less). If SN 18(2) is agent-less, VEM 22(2) may cache the information included in the NSH in a flow table. If SN 18(2) is agent-full, the packet may be forwarded to the agent as appropriate.

According to various embodiments, after the packet is suitably serviced at SN 18(2), VEM 22(2) may intercept the packet and lookup the next service hop. The NSH may be updated to indicate the next service hop as SN 18(4) (rather than WL 20(2), for example). The packet may be forwarded on service overlay 26 to the next service hop, where VEM 22(3) may intercept the packet, and proceed appropriately.

Embodiments of communication system 10 may decentralize the service forwarding decisions, with each VEM 22 making appropriate next service hop decisions. Embodiments of communication system may facilitate termination or re-origination (as appropriate) of service overlay 26, with seamless integration of agent-full and agent-less services. Any kind of network (e.g., enterprise, service provider, etc.) may implement embodiments of communication system 10 as appropriate.

Further, the service forwarding decision at any of VEMs 22(1)-22(3) may be limited to the next-hop of the service chain, rather than all hops of the service chain. For example, the next service hop decision at the classifier VEM (e.g., VEM 22(1)) may determine the first SN (e.g., SN 18(2)) in the service chain and may send the traffic on service overlay 26 to the first SN (e.g., SN 18(2)). The NSH may be written to indicate the source as WL 20(2) and next service hop as SN 18(2): <overlay: source=WL2, destination=SN2>. The service VEM (e.g., VEM 22(2)) at SN 18(2) may simply allow the traffic on service overlay 26 to pass through to SN 18(2).

After the service is delivered at the SN (e.g., SN 18(2)), the SN (e.g., SN 18(2)) may simply send the serviced traffic back on service overlay 26 to where traffic came from (e.g., WL 20(2), or VEM 22(1)). For example, SN 18(2) may write the NSH to indicate the source as SN 18(2) and destination as VEM 22(1): <overlay: source=SN2, destination=VEM1>. The return traffic may be intercepted by the service VEM (e.g., VEM 22(2)) next (or closest) to the SN (e.g., SN 18(2)). The intercepting service VEM (e.g., VEM 22(2)) may make the service forwarding decision, determining the next SN (e.g., SN 18(4)) in the service chain and re-originating the NSH to the next SN (e.g., SN 18(4)). The NSH may be rewritten to indicate the source as VEM 22(2) and destination as SN 18(4): <overlay: source=VEM2, destination=SN4>.

The process of service forwarding can continue from VEMs 22 to SNs 18 until all SNs in the service chain deliver services. The forwarding decision may be based on the pres-

ence or absence of an agent at SN 18. For example, assume merely for example purposes and not as a limitation, that SN 18(4) is agent-less, VEM 22(3) may notice that NSH indicates a destination of SN 18(4), which is agent-less. VEM 22(3) may terminate service overlay 26 and perform translation to send the traffic to SN 18(4). After SN 18(4) delivers the service, it may simply send the original payload packet out, which may be received by VEM 22(3) for translation back onto service overlay 26. VEM 22(3) may intercept SN 18(4)'s traffic and determine the next service hop as SN 18(5) (which, for example purposes, may be agent-full and on the same VEM as SN 18(4)). VEM 22(3) may re-originate NSH to SN 18(5): <overlay: source=VEM3, destination=SN5>. After the service is applied, SN 18(5) may simply re-originate the NSH back to VEM 22(3): <overlay: source=SN5, destination=VEM3>.

The service VEM (e.g., VEM 22(3)) intercepting the return traffic from the last SN (e.g., SN 18(5)) in the service chain may determine the end of service chain. If the last VEM (e.g., VEM 22(3)) is capable of forwarding the payload traffic, it may simply forward it on the underlay network (e.g., network 12). If on the other hand, the payload traffic can only be forwarded by classifier VEM (e.g., VEM 22(1)), the NSH may be re-originated by the last VEM (e.g., VEM 22(3)) back to the classifier VEM (e.g., VEM 22(1)). VEM 22(1) may receive the serviced packet on service overlay 26 and may determine that all services on the service chain are delivered. VEM 22(1) may forward the original payload packet, serviced by the service chain, natively or on the underlay network (e.g., network 12), as appropriate.

In some embodiments, for example, as in a service provider network environment that represents a non-homogeneous environment, the network infrastructure, including DVS 14 may be owned and operated by the provider; WLs 20 may belong to the tenants of the provider; and SNs 18 may be hosted by the provider on behalf of the tenant or hosted by the tenants themselves, or by other third parties. In some embodiments, for example, wherein the service provider hosts SNs 18 on behalf of the tenant, NSH of service overlay 26 may use the IP/MAC addresses of VEMs 22 and SNs 18 for source and destination addresses.

In some other embodiments, for example wherein the tenants host SNs 18, traffic over service overlay 26 may be forwarded in two hops: (1) provider overlay; and (2) tenant overlay. In provider overlay, origination and termination of service overlay 26 may be implemented within the network infrastructure, including DVS 14 (and associated VEMs 22). Hence, the end points of the provider overlay may comprise VEMs 22. For example, turning to the example of service chain 1, one of the provider overlay hops may stretch from VEM 22(1) to VEM 22(2): <overlay: source=VEM1, destination=VEM2>. In tenant overlay, the origination may occur in the network infrastructure and the termination may be a local destination. Continuing the example of service chain 1, one of the tenant overlay hops may stretch from VEM 22(2) to SN 18(2): <overlay: source=VEM2, destination=SN2>. The tenant overlay source address can be local to respective VEMs 22 (e.g., VEM 22(2)), allowing for overcoming the tenant and provider address domain packet forwarding issues across the service provider network.

Embodiments of communication system 10 can provide a method to decentralize the service forwarding decisions across the network infrastructure, enabling overlays to be re-originated within the network infrastructure while allowing the agent-full service nodes to be agnostic to service chaining. Embodiments of communication system 10 can enable service chains that can have agent-full as well as

agent-less services as part of the same service chain. Embodiments of communication system **10** can also provide a method for maintaining the provider and tenant address space separation while allowing the service chains to cross the provider-tenant network boundary.

According to various embodiments of communication system **10**, service chains can be realized in a distributed fashion across the network infrastructure without a centralized bottleneck while keeping the agent-full services agnostic to service forwarding decisions. Service chains may contain both agent-full and agent-less services while being agnostic to service forwarding decisions. Service chaining performance may stay constant and may not degrade with the number of services in the service chain. Service chains may involve services that are hosted by either the provider or the tenant without departing from the scope of the embodiments, providing a 'clean' service chaining implementation in both enterprise and service provider environments (among other network environments).

Turning to the infrastructure of communication system **10**, the network topology can include any number of servers, virtual machines, switches (including distributed virtual switches), routers, and other nodes inter-connected to form a large and complex network. A node may be any electronic device, client, server, peer, service, application, or other object capable of sending, receiving, or forwarding information over communications channels in a network. Elements of FIG. **1** may be coupled to one another through one or more interfaces employing any suitable connection (wired or wireless), which provides a viable pathway for electronic communications. Additionally, any one or more of these elements may be combined or removed from the architecture based on particular configuration needs. Communication system **10** may include a configuration capable of TCP/IP communications for the electronic transmission or reception of data packets in a network. Communication system **10** may also operate in conjunction with a User Datagram Protocol/Internet Protocol (UDP/IP) or any other suitable protocol, where appropriate and based on particular needs. In addition, gateways, routers, switches, and any other suitable nodes (physical or virtual) may be used to facilitate electronic communication between various nodes in the network.

Note that the numerical and letter designations assigned to the elements of FIG. **1** do not connote any type of hierarchy; the designations are arbitrary and have been used for purposes of teaching only. Such designations should not be construed in any way to limit their capabilities, functionalities, or applications in the potential environments that may benefit from the features of communication system **10**. It should be understood that communication system **10** shown in FIG. **1** is simplified for ease of illustration.

The example network environment may be configured over a physical infrastructure that may include one or more networks and, further, may be configured in any form including, but not limited to, local area networks (LANs), wireless local area networks (WLANs), VLANs, metropolitan area networks (MANs), wide area networks (WANs), VPNs, Intranet, Extranet, any other appropriate architecture or system, or any combination thereof that facilitates communications in a network. In some embodiments, a communication link may represent any electronic link supporting a LAN environment such as, for example, cable, Ethernet, wireless technologies (e.g., IEEE 802.11x), ATM, fiber optics, etc. or any suitable combination thereof. In other embodiments, communication links may represent a remote connection through any appropriate medium (e.g., digital subscriber lines (DSL), telephone lines, T1 lines, T3 lines, wireless, satellite, fiber optics, cable,

Ethernet, etc. or any combination thereof) and/or through any additional networks such as a wide area networks (e.g., the Internet).

In various embodiments, services nodes **18(1)-18(5)** represent a specific functionality (e.g., provision of a specific service) and may be embodied in one or more physical appliances. For example, some services nodes (e.g., service nodes **18(4)** and **18(5)**) may be provided in a common network element, whereas some other service nodes (e.g., **18(1)** and **18(2)**) may be stand-alone network elements that are configured to exclusively provide the respective specific service. Note that although only five service nodes **18(1)-18(5)** are illustrated in FIG. **1**, any number of service nodes and corresponding services may be provided within the broad scope of the embodiments.

In various embodiments, workload **20** may be separate computing devices running applications (e.g., server/client applications in client-server network architecture). In other embodiments, workload **20** may be separate virtual machines on the same or different computing devices (e.g., server blades in a data center). In some embodiments, workload **20** may include server blades configured in one or more chassis. DVS **14** may include physical and virtual switches and can include any suitable network element capable of receiving packets, and forwarding packets appropriately in a network environment. Any number of workload may be active within network **12** within the broad scope of the embodiments.

VEMs **20** can include virtual interfaces (e.g., virtual equivalent of physical network access ports) that maintain network configuration attributes, security, and statistics across mobility events, and may be dynamically provisioned within virtualized networks based on network policies stored in DVS **14** as a result of VM provisioning operations by a hypervisor management layer. VEMs **22** may follow virtual network interface cards (vNICs) when VMs move from one physical server to another. The movement can be performed while maintaining port configuration and state, including NetFlow, port statistics, and any Switched Port Analyzer (SPAN) session. By virtualizing the network access port with DPs **24(2)-24(6)**, transparent mobility of VMs across different physical servers and different physical access-layer switches within an enterprise network may be possible. vPaths **24(1)-24(3)** may provide intelligent traffic steering (e.g., flow classification and redirection), and fast path offload for policy enforcement of flows. vPaths **24(1)-24(3)** may be configured for multi-tenancy, providing traffic steering and fast path offload on a per-tenant basis. Although only three vPaths **24(1)-24(3)** are illustrated in FIG. **1**, any number of vPaths may be provided within the broad scope of the embodiments of communication system **10**.

In one example embodiment, service controller **16** may be an application coupled with a management module (e.g., virtual supervisor module (VSM)) of DVS **14**. In another embodiment, service controller **16** may be a stand-alone application (e.g., provisioned in a suitable network element) separate and distinct from DVS **14** and communicating therewith through appropriate communication links. In some embodiments, service controller **16** may be provisioned in the same local area network as workload **20**. In other embodiments, service controller **16** may be provisioned in a different local area network separate and remote from workload **20**. Service controller **16** may include a graphical user interface (GUI) based controller, or a CLI based controller, or a combination thereof.

Turning to FIG. **2**, FIG. **2** is a simplified block diagram illustrating example details that may be associated with an embodiment of communication system **10**. An example ser-

vice chain is illustrated in the figure, starting at WL 20(2), proceeding to SN 18(2), then to SN 18(3), then to SN 18(4), then to SN 18(5), and lastly, to WL 20(5): WL2→SN2→SN3→SN4→SN5→WL5. Service controller 16 may program service forwarding tables 30(1)-30(3) at respective VEMs 22(1)-22(3). Each service forwarding table 30(1)-30(3) may include a path identifier (ID) and a next service hop information. Some SNs 18 may include an agent 32. For example, SN 18(3) may include agent 32. Note that the configuration described herein is merely for example purposes, and is not intended to be a limitation of embodiments of communication system 10.

The packet from WL 20(2) may be encapsulated with the NSH at classifier VEM 22(1) based on information in service forwarding table 30(1). The packet may be forwarded on service overlay 26 to the next service hop, namely SN 18(2). VEM 22(2) may decapsulate the NSH, and forward the packet through interface 34(1) to SN 18(2), which may be agent-less. SN 18(2) may service the packet, and rewrite the packet header to indicate the destination address of VEM 22(1) and send the packet out through interface 34(2). VEM 22(2) may intercept the packet, and re-originate the NSH based on information in service forwarding table 30(2). The destination is written to be the IP/MAC address of SN 18(3). VEM 22(2) may forward the packet to SN 18(3) transparently (e.g., without decapsulation of the NSH) via interface 34(3) as SN 18(3) includes agent 32. After being serviced, the packet may be returned to VEM 22(2) via interface 34(3). VEM 22(2) may intercept the packet, and re-originate the NSH based on information in service forwarding table 30(2). The destination may be written to be the IP/MAC address of SN 18(4) and the packet forwarded to VEM 22(3) on service overlay 26.

VEM 22(3) may decapsulate the packet, and forward the packet to SN 18(4) over interface 34(4). SN 18(4) may service the packet appropriately, and attempt to return it to VEM 22(1) over interface 34(5). VEM 22(3) may intercept the packet, and re-originate the NSH based on information in service forwarding table 30(3). The destination may be written to be the IP/MAC address of SN 18(5) and the packet forwarded to SN 18(5) over interface 34(6). SN 18(5) may service the packet appropriately, and attempt to return it to VEM 22(1) over interface 34(7). VEM 22(3) may intercept the packet, and re-originate the NSH based on information in service forwarding table 30(3). In some embodiments, the destination may be written to be the IP/MAC address of WL 20(5) and the packet forwarded to WL 20(5) over network 12, or the appropriate interface. In other embodiments, the destination may be written to be the IP/MAC address of classifier VEM 22(1) and the packet forwarded to WL 20(2) on service overlay 26 as appropriate.

Turning to FIG. 3, FIG. 3 is a simplified block diagram illustrating example details that may be associated with an embodiment of communication system 10. Data traffic may be communicated on underlay network 12, and service traffic (e.g., packets subject to being serviced at various SNs 18) may be communicated on service overlay 26. An example data traffic path may comprise communication between WL 20(1) and WL 20(4), and the service traffic path may comprise communication from WL 20(1) to SN 18(1), to SN 18(2), to SN 18(3), and back to WL 20(1): WL1→SN1→SN2→SN3→WL1. Assume, merely for example purposes and not as a limitation that SN 18(1) and SN 18(3) are agent-full, including respective agents 32(1) and 32(2), and SN 18(2) is agent-less.

Service controller 16 may configure multiple paths (e.g., a forward path and a return path) on each service forwarding

table 30(1)-30(3). Each path may indicate the service path identifier that indicates (e.g., identifies) the specific service chain, a location on the service chain, indicated by a service index, and a next service node on the service chain. For example, at service forwarding table 30(1) on VEM 22(1), a first path may be configured as P1: seq5: SN1, indicating service chain identified by service path identifier P1, with a service index seq5 corresponding to the location of the packet on service chain P1, and a next service node of SN 18(1) (e.g., when the packet arrives at VEM 22(1) on service chain P1 with service index seq5, send it to SN 18(1)). A second path at service forwarding table 30(1) on VEM 22(1) may be configured as P1: seq1: end, indicating the service chain identified by service path identifier P1, with a service index seq1 corresponding to the end of the return service traffic path at VEM 22(1) (e.g., when the packet arrives at VEM 22(1) on service chain P1 with service index seq1, end the service path).

At service forwarding table 30(2) on VEM 22(2), the first path may be configured as P1: seq5: SN1, indicating the service chain identified by service path identifier P1, with a service index seq5 corresponding to the location seq5 of the packet on the service chain and next service node of SN 18(1) (e.g., when the packet arrives at VEM 22(2) on service chain P1 with service index seq5, send it to SN 18(1)). The second path at service forwarding table 30(2) on VEM 22(2) may be configured as P1: seq4: SN2, indicating the service chain identified by service path identifier P1, with a service index seq4 corresponding to the location seq4 of the packet on the service chain and the next service node of SN 18(2) (e.g., when the packet arrives at VEM 22(2) on service chain P1 with service index seq4, send it to SN 18(2)).

At service forwarding table 30(3) on VEM 22(3), the first path may be configured as P1: seq4: SN2, indicating the service chain identified by service path identifier P1, with a service index seq4 corresponding to the location seq4 of the packet on the service chain and the next service node of SN 18(2) (e.g., when the packet arrives at VEM 22(3) on service chain P1 with service index seq4, send it to SN 18(2)). The second path at service forwarding table 30(3) on VEM 22(3) may be configured as P1: seq3: SN3, indicating the service chain identified by service path identifier P1, with a service index seq3 corresponding to the location seq3 of the packet on the service chain and the next service node of SN 18(3) (e.g., when the packet arrives at VEM 22(3) on service chain P1 with service index seq3, send it to SN 18(3)). The third path at service forwarding table 30(3) on VEM 22(3) may be configured as P1: seq1: VEM1, indicating the service chain identified by service path identifier P1, with a service index seq1 corresponding to the location seq1 of the packet on the service chain and the next service node of VEM 22(1) (e.g., when the packet arrives at VEM 22(3) on path P1 with service index seq1, send it to VEM 22(1)).

Turning to FIG. 4, FIG. 4 is a simplified block diagram illustrating details of an example VEM 22 according to an embodiment of communication system 10. An incoming packet 36 may include NSH 38, a transport header 40, and a payload 42. A receive module 44 at VEM 22 may receive packet 36. A decapsulate module 46 may decapsulate the transport header 40 and NSH 38 and inspect the contents thereof. An encapsulate module 48 may lookup service forwarding table 30 and determine the next service hop to send packet 36. Service forwarding table 30 may be programmed by service controller 16 with appropriate entries.

If packet 36 is to be sent to an agent-full service node 18, for example, which includes agent 32, encapsulate module 50 may encapsulate packet 36 with another NSH 38, and retain

11

transport header 40 and payload 42 to generate outgoing packet 56. If packet 36 is to be sent to an agent-less service node 18, encapsulate module 50 may remove transport header 40 and generate outgoing packet 56 without NSH 38. A transmit module 50 may transmit outgoing packet 56 to appropriate SN 18. A processor 52 and a memory element 54 may facilitate the operations described herein.

Turning to FIG. 5, FIG. 5 is a simplified block diagram illustrating an example service forwarding table entry 57 according to an embodiment of communication system 10. Example service forwarding table entry 57 may include a field 58 comprising the service path ID and a service index. Field 59 may include the next service node on the path identified by the service path ID and the service index of field 58.

Turning to FIG. 6, FIG. 6 is a simplified block diagram illustrating an example service overlay packet format according to an embodiment of communication system 10. Example service overlay packet format may include NSH 38, transport header 40 and payload 42. NSH 38 may comprise four 32-bit context headers (e.g., service shared context, service platform context, network shared context, and network platform context), and an additional header 60 comprising the 24 bit service path identifier and 8 bit service index.

Turning to FIG. 7, FIG. 7 is a simplified flow diagram illustrating example operations 70 that may be associated with an embodiment of communication system 10. At 72, service controller 16 may discover locations of service nodes 18 in DVS 14. At 74, when service chains are instantiated at classifier VEMs 22 (e.g., at WLs 20), service controller 16 may be notified of the service chain instantiation. At 76, service controller 16 may assign respective path identifiers to each service chain. At 78, service controller 16 may populate service forwarding tables 30 with appropriate entries. At 80, service controller 16 may program (e.g., install, implement, execute, etc.) service forwarding tables 30 at respective VEMs 22 based on service node discovery information.

Turning to FIG. 8, FIG. 8 is a simplified flow diagram illustrating example operations 90 that may be associated with an embodiment of communication system 10. At 92, a determination may be made at VEM 22 whether incoming traffic is on service overlay 26 (for example, the determination may be informed by the presence or absence of NSH 38). If traffic is on service overlay 26, at 92, the path identifier and service index information may be obtained from NSH 38. At 96, service forwarding table 30 may be consulted to determine next service node 18 or tuple <path-identifier, service index>. At 98, a determination may be made whether next service node 18 is local to VEM 22. If next service node 18 is local to VEM 22, a determination may be made at 100 whether service node 18 is agent-less. If service node 18 is agent-less, at 102, traffic may be forwarded natively (e.g., without NSH 38) with service header metadata mapped to appropriate VLAN tags. If service node 18 is agent-full, at 104, another NSH 38 may be imposed with appropriate path identifier and service index as determined from service forwarding table 30. At 106, transport header 40 with appropriate source and destination may be imposed to forward packet to next service node 18 and the packet/frame may be appropriately forwarded. Turning back to 98, if next service node 18 is not local to VEM 22, the operations may step to 104, and continue thereafter.

Turning back to 92, if the traffic is not on overlay 92, a determination may be made at 108 whether the traffic is from an agent-less service node. If not, indicating that the traffic is from WL 20, at 110, the traffic may be appropriately classified and the service path may be determined (e.g., service path identifier and service index identified). The operations may

12

step to 96 and continue thereafter. On the other hand, at 92, if the traffic is from an agent-less service node, at 112, the flow may be looked up in appropriate tables and the service path identifier and service index may be determined. The operations may step to 96, and continue thereafter.

Note that in this Specification, references to various features (e.g., elements, structures, modules, components, steps, operations, characteristics, etc.) included in “one embodiment”, “example embodiment”, “an embodiment”, “another embodiment”, “some embodiments”, “various embodiments”, “other embodiments”, “alternative embodiment”, and the like are intended to mean that any such features are included in one or more embodiments of the present disclosure, but may or may not necessarily be combined in the same embodiments. Note also that an ‘application’ as used herein this Specification, can be inclusive of an executable file comprising instructions that can be understood and processed on a computer, and may further include library modules loaded during execution, object files, system files, hardware logic, software logic, or any other executable modules.

In example implementations, at least some portions of the activities outlined herein may be implemented in software in, for example, DVS 14. In some embodiments, one or more of these features may be implemented in hardware, provided external to these elements, or consolidated in any appropriate manner to achieve the intended functionality. The various network elements (e.g., DVS 14, service controller 16) may include software (or reciprocating software) that can coordinate in order to achieve the operations as outlined herein. In still other embodiments, these elements may include any suitable algorithms, hardware, software, components, modules, interfaces, or objects that facilitate the operations thereof.

Furthermore, DVS 14 described and shown herein (and/or their associated structures) may also include suitable interfaces for receiving, transmitting, and/or otherwise communicating data or information in a network environment. Additionally, some of the processors and memory elements associated with the various nodes may be removed, or otherwise consolidated such that a single processor and a single memory element are responsible for certain activities. In a general sense, the arrangements depicted in the FIGURES may be more logical in their representations, whereas a physical architecture may include various permutations, combinations, and/or hybrids of these elements. It is imperative to note that countless possible design configurations can be used to achieve the operational objectives outlined here. Accordingly, the associated infrastructure has a myriad of substitute arrangements, design choices, device possibilities, hardware configurations, software implementations, equipment options, etc.

In some of example embodiments, one or more memory elements (e.g., memory element 54) can store data used for the operations described herein. This includes the memory element being able to store instructions (e.g., software, logic, code, etc.) in non-transitory media, such that the instructions are executed to carry out the activities described in this Specification. A processor can execute any type of instructions associated with the data to achieve the operations detailed herein in this Specification. In one example, processors (e.g., processor 52) could transform an element or an article (e.g., data) from one state or thing to another state or thing. In another example, the activities outlined herein may be implemented with fixed logic or programmable logic (e.g., software/computer instructions executed by a processor) and the elements identified herein could be some type of a programmable processor, programmable digital logic (e.g., a field

13

programmable gate array (FPGA), an erasable programmable read only memory (EPROM), an electrically erasable programmable read only memory (EEPROM)), an ASIC that includes digital logic, software, code, electronic instructions, flash memory, optical disks, CD-ROMs, DVD ROMs, magnetic or optical cards, other types of machine-readable mediums suitable for storing electronic instructions, or any suitable combination thereof.

These devices may further keep information in any suitable type of non-transitory storage medium (e.g., random access memory (RAM), read only memory (ROM), field programmable gate array (FPGA), erasable programmable read only memory (EPROM), electrically erasable programmable ROM (EEPROM), etc.), software, hardware, or in any other suitable component, device, element, or object where appropriate and based on particular needs. The information being tracked, sent, received, or stored in communication system 10 could be provided in any database, register, table, cache, queue, control list, or storage structure, based on particular needs and implementations, all of which could be referenced in any suitable timeframe. Any of the memory items discussed herein should be construed as being encompassed within the broad term 'memory element.' Similarly, any of the potential processing elements, modules, and machines described in this Specification should be construed as being encompassed within the broad term 'processor.'

It is also important to note that the operations and steps described with reference to the preceding FIGURES illustrate only some of the possible scenarios that may be executed by, or within, the system. Some of these operations may be deleted or removed where appropriate, or these steps may be modified or changed considerably without departing from the scope of the discussed concepts. In addition, the timing of these operations may be altered considerably and still achieve the results taught in this disclosure. The preceding operational flows have been offered for purposes of example and discussion. Substantial flexibility is provided by the system in that any suitable arrangements, chronologies, configurations, and timing mechanisms may be provided without departing from the teachings of the discussed concepts.

Although the present disclosure has been described in detail with reference to particular arrangements and configurations, these example configurations and arrangements may be changed significantly without departing from the scope of the present disclosure. For example, although the present disclosure has been described with reference to particular communication exchanges involving certain network access and protocols, communication system 10 may be applicable to other exchanges or routing protocols. Moreover, although communication system 10 has been illustrated with reference to particular elements and operations that facilitate the communication process, these elements, and operations may be replaced by any suitable architecture or process that achieves the intended functionality of communication system 10.

Numerous other changes, substitutions, variations, alterations, and modifications may be ascertained to one skilled in the art and it is intended that the present disclosure encompass all such changes, substitutions, variations, alterations, and modifications as falling within the scope of the appended claims. In order to assist the United States Patent and Trademark Office (USPTO) and, additionally, any readers of any patent issued on this application in interpreting the claims appended hereto, Applicant wishes to note that the Applicant: (a) does not intend any of the appended claims to invoke paragraph six (6) of 35 U.S.C. section 112 as it exists on the date of the filing hereof unless the words "means for" or "step for" are specifically used in the particular claims; and (b) does

14

not intend, by any statement in the specification, to limit this disclosure in any way that is not otherwise reflected in the appended claims.

What is claimed is:

1. A method, comprising:

receiving a packet belonging to a service chain in a distributed virtual switch (DVS) network environment, wherein the packet includes a network service header (NSH) indicating a service path identifier identifying the service chain and a location of the packet on the service chain;

evaluating a service forwarding table configured by a service controller for a next service node based on the service path identifier and the location, wherein a plurality of different forwarding tables are distributed across the DVS at a corresponding plurality of virtual Ethernet Modules (VEMs) associated with respective service nodes in the service chain, wherein the service controller configures the service forwarding table at a first VEM when the service chain is instantiated at a second VEM, wherein the first VEM makes an independent forwarding decision based at least on the service forwarding table at the first VEM, and the service path identifier and the location indicated in the NSH of the packet; and

forwarding the packet to the next service node, wherein substantially all services in the service chain are provided sequentially to the packet in a single service loop on a service overlay.

2. The method of claim 1, wherein the second VEM comprises a classifier VEM attached to a workload that forms an initial node of the service chain.

3. The method of claim 1, wherein the service controller discovers the service nodes associated with the service chain, wherein the service controller configures the service forwarding tables at the corresponding VEMs associated with the respective service nodes.

4. The method of claim 1, wherein the service controller assigns the service path identifier to the service chain.

5. The method of claim 1, wherein if the next service node does not include an agent, the packet is forwarded without the NSH.

6. The method of claim 1, wherein if the next service node includes an agent, the packet is forwarded with the NSH.

7. The method of claim 1, wherein if the packet is received from an agent-less service node, the service forwarding table is looked up to determine the service path identifier and the location.

8. The method of claim 1 wherein the service path identifier is a 24 bit value, and the location is specified by an 8 bit long service index.

9. Non-transitory tangible media that includes instructions for execution, which when executed by a processor, is operable to perform operations comprising:

receiving a packet belonging to a service chain in a DVS network environment, wherein the packet includes a NSH indicating a service path identifier identifying the service chain and a location of the packet on the service chain;

evaluating a service forwarding table configured by a service controller for a next service node based on the service path identifier and the location, wherein a plurality of different forwarding tables are distributed across the DVS at a corresponding plurality of VEMs associated with respective service nodes in the service chain, wherein the service controller configures the service forwarding table at a first VEM when the service

15

chain is instantiated at a second VEM, wherein the first VEM makes an independent forwarding decision based at least on the service forwarding table at the first VEM, and the service path identifier and the location indicated in the NSH of the packet; and

forwarding the packet to the next service node, wherein substantially all services in the service chain are provided sequentially to the packet in a single service loop on a service overlay.

10. The media of claim 9, wherein the service controller assigns the service path identifier to the service chain.

11. The media of claim 9, wherein the service path identifier is a 24 bit value, and the location is specified by an 8 bit long service index.

12. The media of claim 9, wherein the service controller discovers the service nodes associated with the service chain, wherein the service controller configures the service forwarding tables at the corresponding VEMs associated with the respective service nodes.

13. The media of claim 9, wherein if the next service node does not include an agent, the packet is forwarded without the NSH.

14. The media of claim 9, wherein if the next service node includes an agent, the packet is forwarded with the NSH.

15. An apparatus, comprising:

a VEM in a DVS network environment, wherein the VEM includes a memory element for storing data, and a processor, wherein the processor executes instructions associated with the data, wherein the processor and the memory element cooperate, such that the apparatus is configured for:

receiving a packet belonging to a service chain in a DVS network environment, wherein the packet includes a NSH indicating a service path identifier identifying the service chain and a location of the packet on the service chain;

16

evaluating a service forwarding table configured by a service controller for a next service node based on the service path identifier and the location, wherein a plurality of different forwarding tables are distributed across the DVS at a corresponding plurality of VEMs associated with respective service nodes in the service chain, wherein the service controller configures the service forwarding table at a first VEM when the service chain is instantiated at a second VEM, wherein the first VEM makes an independent forwarding decision based at least on the service forwarding table at the first VEM, and the service path identifier and the location indicated in the NSH of the packet; and

forwarding the packet to the next service node, wherein substantially all services in the service chain are provided sequentially to the packet in a single service loop on a service overlay.

16. The apparatus of claim 15, wherein the service controller assigns the service path identifier to the service chain.

17. The apparatus of claim 15, wherein the service path identifier is a 24 bit value, and the location is specified by an 8 bit long service index.

18. The apparatus of claim 15, wherein the service controller discovers the service nodes associated with the service chain, wherein the service controller configures the service forwarding tables at the corresponding VEMs associated with the respective service nodes.

19. The apparatus of claim 15, wherein if the next service node does not include an agent, the packet is forwarded without the NSH.

20. The apparatus of claim 15, wherein if the next service node includes an agent, the packet is forwarded with the NSH.

* * * * *